



(19)

(11) Publication number:

1

Generated Document.

PATENT ABSTRACTS OF JAPAN(21) Application number: **10024592**(51) Intl. Cl.: **G06F 3/06 G06F 3/06 G06F 12/16**(22) Application date: **05.02.98**

(30) Priority:	(71) Applicant: TOSHIBA CORP
(43) Date of application publication: 17.08.99	(72) Inventor: MOTOSAWA KUNIAKI
(84) Designated contracting states:	(74) Representative:

**(54) FAULT AVOIDING
METHOD FOR STORAGE
AREA IN LOG IN TYPE
DISK STORAGE DEVICE,
DISK STORAGE DEVICE,
AND COMPUTER-
READABLE STORAGE
MEDIUM USED IN THE
SAME DEVICE**

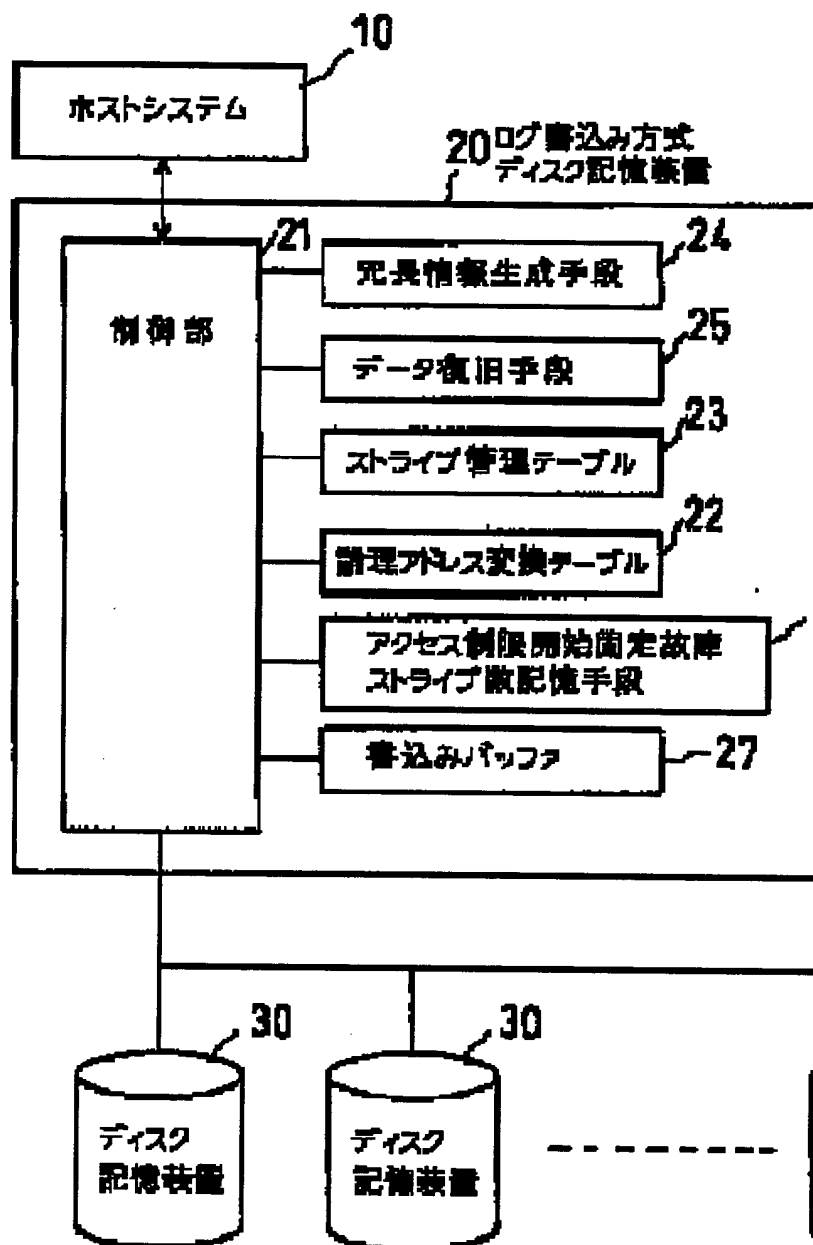
(57) Abstract:

PROBLEM TO BE SOLVED: To provide a highly reliable, high speed log in type disk storage device by improving avoiding of a fault due to a fixed failure in a part of storage area, a fault avoiding method in the storage area in the log in type disk storage device, and a computer-readable storage medium where program information of the same method is stored.

SOLUTION: A log in type disk storage 20 which connects a host system 10 and a disk storage 30 is provided internally with a control part

21 and a logical address conversion table 22, a stripe management table 23, a redundant information generating means 24, a data recovering means 25, an access limitation start fixed fault stripe number storage means 26, and a write buffer 27. Consequently, the avoiding of a fault due to a fixed failure in a part of storage area can be improved and highly reliable, high speed performance can be realized.

COPYRIGHT: (C)1999,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-224166

(43) 公開日 平成11年(1999) 8月17日

(51) Int.Cl.⁸

G 0 6 F 3/06

12/16

識別記号

5 4 0

3 0 5

3 1 0

3 2 0

F I

G 0 6 F 3/06

12/16

5 4 0

3 0 5 F

3 1 0 Q

3 2 0 L

審査請求 未請求 請求項の数6 OL (全 13 頁)

(21) 出願番号 特願平10-24592

(22) 出願日 平成10年(1998) 2月5日

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 本沢 邦朗

東京都青梅市末広町2丁目9番地 株式会
社東芝青梅工場内

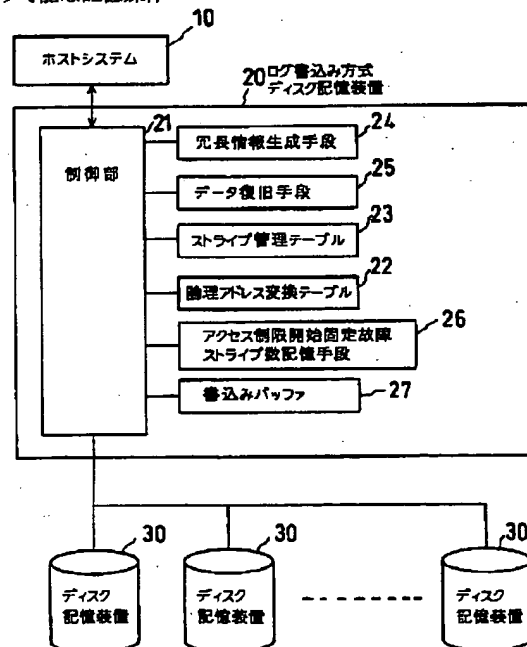
(74) 代理人 弁理士 大胡 典夫 (外1名)

(54) 【発明の名称】 ログ書き込み方式のディスク記憶装置に於ける記憶領域の障害回避方法、及びディスク記憶装置、同装置に用いられるコンピュータ読込み可能な記憶媒体

(57) 【要約】

【課題】 一部の記憶領域の固定故障による障害回避の改善を図り、高信頼且つ高速なログ書き込み方式に於けるディスク記憶装置、及びログ書き込み方式のディスク記憶装置に於ける記憶領域の障害回避方法、同方法のプログラム情報を格納したコンピュータ読取り可能な記憶媒体を提供すること。

【解決手段】 ホストシステム10及びディスク記憶装置30を接続するログ書き込み方式ディスク記憶装置20は、制御部21と論理アドレス変換テーブル22、ストライプ管理テーブル23、冗長情報生成手段24、データ復旧手段25、アクセス制限開始固定故障ストライプ数記憶手段26、及び書き込みバッファ27を内設する。これにより、一部の記憶領域の固定故障による障害回避の改善を図れ、高信頼且つ高速性能を実現できる。



【特許請求の範囲】

【請求項1】 複数のディスク記憶装置を有し、耐故障性を実現する冗長情報を生成する冗長情報生成手段と、故障が発生した際に上記冗長情報生成手段の冗長情報を用いてデータを復旧するデータ復旧手段と、ストライプ中に格納されている有効ブロック数を保持するストライプ管理テーブル手段と、論理アドレスに対応するストライプ番号と上記ストライプ管理テーブル手段のストライプ中の位置を示すブロック番号を保持する論理アドレス変換テーブル手段とを備えたログ書き込み方式記憶装置に於いて、上記ストライプ管理テーブル手段に有効ブロック数に加え、固定故障セクタを包含するブロックが上記ストライプ中に存在するか否かを示す固定故障セクタ存在フラグを保持させる手段と、あるセクタに固定故障が発生した場合、上記ストライプ管理テーブル手段から空きストライプを検索する検索手段と、上記固定故障セクタを包含するストライプの故障セクタ以外のデータを読み出す読み出し手段とを具備し、この読み出し手段にて読み出したデータを用いて前記データ復旧手段により故障セクタのデータを復旧し、前記読み出し手段にて読み出したデータと復旧したデータからなるストライプを当該空きストライプに書き込むと共に、上記論理アドレス変換テーブル手段の固定故障セクタを含むストライプのストライプ番号を当該空きストライプのストライプ番号に変更し、上記ストライプ管理テーブル手段中の故障セクタを含むストライプの有効ブロック数をストライプ管理テーブル中の当該空きストライプの有効ブロック数にコピーし、上記ストライプ管理テーブル手段中の故障セクタを含むストライプの固定故障セクタ存在フラグをセットすることにより、以降は故障セクタを含むストライプは利用しないことを特徴とするログ書き込み方式に於けるディスク記憶装置。

【請求項2】 アクセス制限開始固定故障ストライプ数記憶手段を設け、上記ストライプ管理テーブル手段中から故障セクタを包含するストライプ数を調べ、固定故障ストライプ数がアクセス制限開始固定故障ストライプ数記憶手段に記憶された個数よりも多くなった場合、書き込みアクセスを禁止することを特徴とする請求項1記載のログ書き込み方式に於けるディスク記憶装置。

【請求項3】 複数のディスク記憶装置を有し、耐故障性を実現する冗長情報を生成する冗長情報生成手段と、故障が発生した際に上記冗長情報生成手段の冗長情報を用いてデータを復旧するデータ復旧手段と、ストライプ中に格納されている有効ブロック数を保持するストライプ管理テーブル手段と、論理アドレスに対応するストライプ番号と上記ストライ

プ管理テーブル手段のストライプ中の位置を示すブロック番号を保持する論理アドレス変換テーブル手段とを備えたログ書き込み方式記憶装置に於いて、上記ストライプ管理テーブル手段に有効ブロック数に加え、固定故障セクタを包含するブロックが上記ストライプ中に存在するか否かを示す固定故障セクタ存在フラグを保持させ、あるセクタに固定故障が発生した場合、上記ストライプ管理テーブル手段から空きストライプを検索し、上記固定故障セクタを包含するストライプの故障セクタ以外のデータを読み出し、この読み出したデータを用いて前記データ復旧手段により故障セクタのデータを復旧し、前記読み出したデータと復旧したデータからなるストライプを当該空きストライプに書き込むと共に、上記論理アドレス変換テーブル手段の固定故障セクタを含むストライプのストライプ番号を当該空きストライプのストライプ番号に変更し、上記ストライプ管理テーブル手段中の故障セクタを含むストライプの有効ブロック数をストライプ管理テーブル中の当該空きストライプの有効ブロック数にコピーし、上記ストライプ管理テーブル手段中の故障セクタを含むストライプの固定故障セクタ存在フラグをセットすることにより、以降は故障セクタを含むストライプは利用しないことを特徴とするログ書き込み方式のディスク記憶装置に於ける記憶領域の障害回避方法。

【請求項4】 上記ストライプ管理テーブル手段中から故障セクタを包含するストライプ数を調べ、固定故障ストライプ数が記憶された個数よりも多くなった場合、書き込みアクセスを禁止することを特徴とする請求項3記載のログ書き込み方式のディスク記憶装置に於ける記憶領域の障害回避方法。

【請求項5】 複数のディスク記憶装置を有し、耐故障性を実現する冗長情報を生成する冗長情報生成手段と、故障が発生した際に上記冗長情報生成手段の冗長情報を用いてデータを復旧するデータ復旧手段と、ストライプ中に格納されている有効ブロック数を保持するストライプ管理テーブル手段と、論理アドレスに対応するストライプ番号と上記ストライプ管理テーブル手段のストライプ中の位置を示すブロック番号を保持する論理アドレス変換テーブル手段とを備えたログ書き込み方式記憶装置に用いられるコンピュータ読取り可能な記憶媒体であって、上記ストライプ管理テーブル手段に有効ブロック数に加え、固定故障セクタを包含するブロックが上記ストライプ中に存在するか否かを示す固定故障セクタ存在フラグを保持させ、あるセクタに固定故障が発生した場合、上記ストライプ管理テーブル手段から空きストライプを検索し、上記固定故障セクタを包含するストライプの故障セクタ

以外のデータを読み出し、この読み出したデータを用いて前記データ復旧手段により故障セクタのデータを復旧し、前記読み出したデータと復旧したデータからなるストライプを当該空きストライプに書き込むと共に、上記論理アドレス変換テーブル手段の固定故障セクタを含むストライプのストライプ番号を当該空きストライプのストライプ番号に変更し、上記ストライプ管理テーブル手段中の故障セクタを含むストライプの有効ブロック数をストライプ管理テーブル中の当該空きストライプの有効ブロック数にコピーし、上記ストライプ管理テーブル手段中の故障セクタを含むストライプの固定故障セクタ存在フラグをセットすることにより、以降は故障セクタを含むストライプは利用しないことを特徴とするログ書き込み方式のディスク記憶装置に於ける記憶領域の障害回避方法のプログラム情報を格納したコンピュータ読取り可能な記憶媒体。

【請求項6】 上記ストライプ管理テーブル手段中から故障セクタを包含するストライプ数を調べ、固定故障ストライプ数が記憶された個数よりも多くなった場合、書き込みアクセスを禁止することを特徴とする請求項5記載のログ書き込み方式のディスク記憶装置に於ける記憶領域の障害回避方法のプログラム情報を格納したコンピュータ読取り可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、高信頼性及び高速性の実現を図った外部記憶装置、特に一部の記憶領域の固定故障による障害回避の改善を図ったログ書き込み方式に於けるディスク記憶装置、及びログ書き込み方式のディスク記憶装置に於ける記憶領域の障害回避方法、同方法のプログラム情報を格納したに関するコンピュータ読取り可能な記憶媒体に関する。

【0002】

【従来の技術】従来、ログ書き込み方式によるディスク記憶装置では、例えば特開平6-214720号「ディスク記憶装置のデータ更新方法」にあるようなログ書き込み方式が提案されていた。ところが、このものにあっては、ログ書き込み方式のディスク記憶装置に於いて、一部の記憶領域が固定故障した場合の処理に関しては何等解決方法は開示されていないものであった。

【0003】そこで、ログ書き込み方式のディスク記憶装置に、一般的なディスク記憶装置及びディスクアレイ記憶装置に於ける処理を適用した場合について説明する。以下では、図14に示すログ書き込み方式ディスク記憶装置に於いて一部の記憶領域が固定故障した際の処理について説明する。

【0004】ログ書き込み方式ディスク記憶装置90はホストシステム100に接続し、制御部91と1台以上のディスク記憶装置92、92、…から構成される。デ

ィスク記憶装置92は、ディスク92aとディスクを制御するドライブコントローラ92bから構成される。ここで、固定故障した領域を含むセクタを故障セクタと呼ぶ。セクタとはディスクへのアクセスの最小単位である。一般に、ディスク記憶装置は、故障セクタの代わりに利用するための代替セクタを一定数有する代替セクタ領域を確保している。

【0005】(1) ディスク記憶装置での処理

一部の記憶領域の固定故障はディスク記憶装置92内で検出され、まず装置内での回避処理が行われる。この回避処理は、故障が書き込み時に検出された場合と、読み出し時に検出された場合とでは処理が異なるので、各々の場合に分けて説明する。

【0006】a) 書き込み時に固定故障を検出

ドライブコントローラ92bは、書き込みで検出された故障セクタへ書き込む代わりに、予め確保しておいた代替セクタに書き込みを行う。また、故障セクタが代替セクタに振り替わったことを記憶しておく。この情報を用いて以降は故障セクタへのアクセスが発生した場合は、代替セクタがアクセスされるようにドライブコントローラ92bにより制御を行う。この場合、ディスク記憶装置92の外部からは、固定故障が発生したことを何等意識する必要はない。

【0007】ところが、代替セクタ領域は一般に固定である。そのため、代替セクタが全て利用されると、それ以降に書き込みアクセス時に一部の記憶領域の固定故障が検出された場合は、ドライブコントローラ92bはログ書き込み方式ディスク記憶装置90の制御部91に対して書き込みエラーを通知する。代替セクタ領域を予め大きく確保することにより、許容できる故障セクタ数を増やすことができる。しかし、通常時に利用できるディスク容量が減ることが問題となる。

【0008】b) 読み出し時に固定故障を検出

この場合は、ディスク記憶装置において回避できないので、ドライブコントローラ92bは、ログ書き込み方式ディスク記憶装置90の制御部91に対して読み出しエラーを通知する。

【0009】ログ書き込み方式記憶装置が複数のディスク記憶装置から構成されており、且つ耐故障性を実現するために冗長情報を持つ場合は、ドライブコントローラ92bがエラーを通知した場合、以下の(2)の処理を行うことにより固定故障による障害を回避できる。ここで、耐故障性を実現する方法としては、例えばログ書き込み方式記憶装置を構成するディスク記憶装置の1台分の記憶領域をパリティ情報を格納するために用いる方法がある。この場合、1台のディスク記憶装置が利用できなくなっても、このパリティ情報と他のディスク記憶装置のデータを用いてデータ復旧が可能である。

【0010】(2) ログ書き込み方式記憶装置での処理
前述(1)に於いてディスクコントローラが通知したエ

ラーが、読み出しで発生した場合と、書き込みで発生した場合により、各々以下のように処理が行われる。

【0011】a) 読み出しエラーが通知された場合
ログ書き込み方式ディスク記憶装置90の制御部91は、読み出しエラーが発生したデータ復元を行うために必要なデータを読み出す。例えば、故障セクタを包含した冗長情報を構成していた他のディスク記憶装置92内のデータと、この冗長情報を読み出す。その後、この情報により故障セクタに記憶されていたデータを復旧する。次に、復旧したデータを故障セクタに上書きする。これは、故障セクタに対して書き込みを行うことになるので、前述(1)のa)の処理が行われる。

【0012】前述(1)のa)処理では、故障セクタの代替セクタへの振り替え処理が行われる。その際、利用可能な代替セクタがあれば故障セクタは代替セクタに振り替えられ、継続して動作できる。ところが、代替セクタが全て利用されていれば、故障セクタへの振り替えができないため、ログ書き込み方式記憶装置において次のb)の書き込みエラー処理が行われる。

【0013】b) 書き込みエラーが通知された場合
まず、ログ書き込み方式ディスク記憶装置90の制御部91は、故障発生したディスク記憶装置92を構成から論理的に切り離して動作を継続する。この場合、耐故障性が失われた状態になる。これは、更にディスク記憶装置92が故障すると、このログ書き込み方式ディスク記憶装置90全体のデータが失われてしまうためである。

【0014】次に、例えばサービスマン等の保守要員が故障発生したディスク記憶装置を取り出し、代替ディスク記憶装置を追加する。その後、ログ書き込み方式ディスク記憶装置90の制御部91により、ログ書き込み方式記憶装置90を論理的に再構成する。尚、再構成処理中のログ書き込み方式ディスク記憶装置90へのアクセスは著しく性能が低下するため、再構成処理の間はログ書き込み方式ディスク記憶装置90へのアクセスを禁止する場合もある。また、代替セクタ領域に含まれる数のセクタしか固定故障していないのにディスク記憶装置を交換することになる。

【0015】

【発明が解決しようとする課題】上述した従来のログ書き込み方式ディスク記憶装置に於いて一部の記憶領域が固定故障した際に行われる処理の問題を以下に示す。

【0016】即ち、ドライブコントローラ92bが書き込みエラーを通知した場合は、ログ書き込み方式ディスク記憶装置90の制御部91は、故障が発生したディスク記憶装置92を構成から論理的に切り離して動作を継続する。この場合、故障が発生したディスク装置を交換するまでの間は耐故障性が失われる。

【0017】その後、故障が発生したディスク記憶装置を新しいディスク記憶装置と交換し、ログ書き込み方式ディスク記憶装置90の制御部91により、ログ書き込

み方式記憶装置90を論理的に再構成する。再構成処理中のログ書き込み方式記憶装置90へのアクセスは著しく性能が低下してしまう。そのため、再構成処理の間は、ログ書き込み方式記憶装置90へのアクセスを禁止しなければならない場合もあった。また、代替セクタ領域に含まれる数のセクタしか固定故障していないのに、ディスク記憶装置を交換することになり、ディスク記憶装置の寿命を短くしていたものであった。

【0018】尚、ディスク記憶装置の代替セクタの数は動的に変更できないので、予め代替セクタの数を多く確保した場合にはディスク記憶装置の利用効率が落ちることになるものである。

【0019】そこで、本発明は上記事情を考慮して成されたもので、上記不具合を解消し、一部の記憶領域の固定故障による障害回避の改善を図り、高信頼且つ高速なログ書き込み方式に於けるディスク記憶装置、及びログ書き込み方式のディスク記憶装置に於ける記憶領域の障害回避方法、同方法のプログラム情報を格納したコンピュータ読取り可能な記憶媒体を提供することを目的とする。

【0020】

【課題を解決するための手段】本発明は上記目的を達成するため、以下の通りとする。

【0021】(1)複数のディスク記憶装置から構成されており、耐故障性を実現するための冗長情報を生成する冗長情報生成手段と、故障が発生した際に冗長情報を用いてデータを復旧するデータ復旧手段と、ストライプ中に格納されている有効ブロック数を保持するストライプ管理テーブルと、論理アドレスに対応するストライプ番号とストライプ中の位置を示すブロック番号を保持する論理アドレス変換テーブルとを備え、ログ書き込み方式記憶装置に於いて、上記ストライプ管理テーブルに有効ブロック数に加え、固定故障セクタを含むブロックがストライプ中に存在するかどうかを示す固定故障セクタ存在フラグを保持させ、あるセクタに固定故障が発生した場合、ストライプ管理テーブルから空きストライプを検索し、固定故障セクタを含むストライプの故障セクタ以外のデータを読み出し、読み出したデータを用いてデータ復旧手段により故障セクタのデータを復旧し、読み出したデータと復旧したデータからなるストライプを当該空きストライプに書き込むと共に論理アドレス変換テーブルの固定故障セクタを含むストライプのストライプ番号を当該空きストライプのストライプ番号に変更し、ストライプ管理テーブル中の故障セクタを含むストライプの有効ブロック数をストライプ管理テーブル中の当該空きストライプの有効ブロック数にコピーし、ストライプ管理テーブル中の故障セクタを含むストライプの固定故障セクタ存在フラグをセットすることにより、以降は故障セクタを含むストライプは利用しないことを特徴とする。

(2)上記(1)に於いて、アクセス制限開始固定故障ストライプ数記憶手段を備え、ストライプ管理テーブル中から故障セクタを含むストライプの数を調べ固定故障ストライプ数がアクセス制限開始固定故障ストライプ数記憶手段に記憶された個数よりも多くなった場合は、書き込みアクセスを禁止することを特徴とする。

【0022】

【発明の実施の形態】以下、図面を参照して本発明の一実施形態を説明する。

【0023】図1は一部の記憶領域の固定故障回避方法を適用したログ書き込み方式記憶装置の概略構成を示すブロック図であり、ホストシステム10に接続される一部の記憶領域の固定故障回避方法を適用したログ書き込み方式ディスク記憶装置20は、例えば8台のディスク記憶装置30、30、…、30を接続する。そして、ログ書き込み方式ディスク記憶装置20は、制御部21と論理アドレス変換テーブル22、ストライプ管理テーブル23、冗長情報生成手段24、データ復旧手段25、アクセス制限開始固定故障ストライプ数記憶手段26、及び書き込みバッファ27を内設する。

【0024】制御部21は、ホストシステム10やディスク記憶装置30と各種制御情報やデータのやりとりを行うと共に、論理アドレス変換テーブル22やストライプ管理テーブル23、冗長情報生成手段、データ復旧手段、アクセス制限開始固定故障ストライプ数記憶手段26並びに書き込みバッファ27に対し各種制御を行い、ログ書き込み方式ディスク記憶装置20全体の制御処理を司るものである。

【0025】論理アドレス変換テーブル22は、図2に示すように、論理アドレスに対応するデータブロックが格納されているストライプ番号とストライプ内での位置を保持する。

【0026】ストライプ管理テーブル23は、図3に示すように、ストライプ中に格納されている有効ブロック数と、固定故障セクタ存在フラグを保持する。有効ブロック数が0の場合は、そのストライプが空きストライプであることを意味する。固定故障セクタ存在フラグは、そのストライプ内に固定故障セクタが含まれている場合にセットされる。そうでない場合はクリアされている。

【0027】図4はストライプとブロックについて示す図であり、ストライプは複数のブロックを包含するものである。また、図4は、ディスク記憶装置の台数と同様に8個のブロックを包含する例である。ブロックは、ある大きさのデータを記憶できる。

【0028】尚、この中の一つのブロックは、ストライプ中の一部のデータが壊れた際にデータを復旧するための冗長情報を格納する。物理的には、各ブロックは複数のディスク記憶装置に分散して配置されている。

【0029】冗長情報生成手段24は、ストライプ中の一部のデータが壊れた際にデータを復旧できるように冗

長情報を生成する。図5に示すように、1ストライプ中のブロック数を8個とし、8台のディスク記憶装置に1個ずつブロックを分散して配置した場合を例として説明する。この場合、1台のディスク記憶装置が利用不可能になった場合にもデータの復旧ができるように、1個のブロックにはパリティ(P)を格納し、1つのストライプ中の7個のブロックがデータ(D0~D6)を格納する。そして、冗長情報生成手段24は、7個のブロックの内容から排他的論理和を計算し、そのパリティを生成する。尚、図5の例では、B7にパリティ情報を配置しているが、ストライプ内のパリティブロックの位置はストライプ毎に変わっていてもよいことは勿論である。

【0030】ところで、図5では、1台のディスク記憶装置が1ブロックを格納しているが、1台のディスク記憶装置にN個のブロックを格納する場合は、1台のディスク記憶装置が利用不可能になった際にもデータの復旧ができるようにするため、N個のブロックをパリティのために利用する必要がある。

【0031】データ復旧手段25は、上述した冗長情報生成手段24によりストライプ中には冗長情報が格納されており、一部のデータが壊れた際にデータ復旧を行う。図5にあっては、同一ブロックの中であれば最大で1ブロック分のデータが壊れても、他のディスク記憶装置に格納されているデータを用いることにより、壊れたデータを復元することができる。図5では、例えばD0の内容が壊れた際には、データD1~D6とパリティPの排他的論理和を計算することによりD0の内容を復元することができるものである。

【0032】固定故障した領域を含むストライプの個数がアクセス制限開始固定故障ストライプ数記憶手段26に記憶されている数を超えた場合は、ログ書き込み方式記憶装置へのアクセスを制限開始する。

【0033】上記構成につき、故障セクタが検出された際の処理を図6乃至図12を参照して説明する。各図においてa)は故障検出時を、b)は障害回避処理後の状態を示すものである。

【0034】図6に示すように、ストライプS0のブロックB3(格納されているデータはD3)が故障したものとす。尚、以下では、故障セクタを含むブロックを故障ブロック、故障セクタを含むストライプを故障ストライプと呼ぶ。この例では、故障ストライプはS0、故障ブロックはB3である。

【0035】ディスク記憶装置からのエラー通知を受けて、ログ書き込み方式記憶装置の制御部21は以下の処理を行う。

【0036】1)ストライプ管理テーブル23から空ストライプを検索する(図7のステップS702)。以降、検索した空きストライプを新ストライプと呼ぶ。図9のa)のストライプ管理テーブル23の状態により、S1は有効ブロック数が0であり未使用であるのでS1

を新ストライプとする。

【0037】2) 故障ストライプ中の故障セクタ以外のデータとパリティを読み出す(ステップS704)。図8のa)に示すD0~D2、D4~D6、Pを読み出す。

【0038】3) データ復旧手段25により、読み出したデータを用いて故障セクタのデータを復元する(ステップS706)。D0~D2、D4~D6、PからD3のデータを復元する。それをD3'とする。

【0039】4) 上記2)で読み出したデータと、復旧した故障セクタのデータを、故障ストライプ中の相対位置を維持したまま新ストライプに書き込む(ステップS708)。図6のb)の状態になる。即ち、

書き込まれるS1内のアドレス		
D0	→	B0
D1	→	B1
D2	→	B2
D3'	→	B3
D4	→	B4
D5	→	B5
D6	→	B6
P	→	B7

5) 論理アドレス変換テーブル22から、故障ストライプのストライプ番号を検索し(ステップS710)、それを新ストライプのストライプ番号に変更する(ステップS712)。論理アドレス変換テーブル22を検索すると、L0~L6のストライプ番号が故障ストライプのストライプ番号S0であるので、それを新ストライプのストライプ番号S1に変更する。その結果、図8のb)の状態になる。

【0040】6) 故障ストライプ番号をキーとしてストライプ管理テーブル23を参照し、有効ブロック数を読み出す(ステップS714)。更に新ストライプの番号をキーとしてストライプ管理テーブル23を参照して、読み出した有効ブロック数を書き込む(ステップS716)。図9のa)参照。S0の有効ブロック数は7。S1の有効ブロック数を7に変更する。

【0041】7) 故障ストライプの番号をキーにしてストライプ管理テーブル23を参照し、固定故障セクタ存在フラグをセットする(ステップS718)。S0の固定故障セクタ存在フラグをセットする。ストライプ管理テーブル23の状態は図8のb)の状態になる。

【0042】次に、通常時の書き込み及び読み出し処理を示す。

【0043】(1) 書き込み処理

書き込み処理の概要を書き込みバッファの機能とともに説明する。書き込みバッファ27は1ストライプと同じ容量を持ち、図5のようにストライプと同様にブロックを配置する。

【0044】1) 制御部21は、ホストシステム10か

ら受け取った書き込みデータを、ブロックサイズ単位で分割し、一旦書き込みバッファ27に蓄積する(図10のステップS1002)。バッファに蓄積する際にブロックの論理アドレスを書き込みバッファ27上のタグの領域に記録する。

【0045】2) 書き込みバッファ27に蓄積されたブロックの数が、1ストライプ中に含まれるブロック数から冗長情報(パリティ)のために必要なブロック数を引いた個数になると、冗長情報用のブロックの領域を確保する(ステップS1004)。そして、蓄積されているブロックの内容を用いて、冗長情報生成手段24により冗長情報を生成する(ステップS1006)。冗長情報は、書き込みバッファ27上に確保した領域に書き込む。

【0046】3) 書き込みバッファ27の内容をディスク上の空ストライプに一括して書き込む(ステップS1008)。ここで、空きストライプとは有効ブロック数が0であり、固定故障存在フラグがセットされていないストライプのことである。即ち、固定故障存在フラグがセットされているストライプは、空きストライプとして書き込みの対象とならない。また、この際、ストライプ管理テーブル23中の有効ブロック数を変更する。これは、書き込みを行った空きストライプの有効ブロック数の更新と、旧データを持つブロックを格納しているストライプの有効ブロック数の更新の2つの処理からなる。

【0047】旧データを持つブロックを格納しているストライプの有効ブロック数の更新は次のように行う。

【0048】まず、書き込みバッファ27に蓄積された各ブロックの論理アドレスをキーとして論理アドレス変換テーブル22を参照し、該当論理アドレスの旧データが格納されているストライプを特定する(図11のステップS1102)。その後、論理アドレス変換テーブル22中の、各ブロックの論理アドレスをキーとしたエントリのストライプ番号には、一括書き込みを行った空きストライプの番号を、ブロック番号には書き込みバッファ27内での位置(=ストライプ内での位置)を書き込む。

【0049】次に、先に特定した旧データを格納するストライプをキーにしてストライプ管理テーブル23を参照し、有効ブロック数の値を1減じる。この処理を、書き込みバッファ27上の全てのデータが格納されているブロックに対して行う(ステップS1104)。

【0050】(2) 読み出し処理

続いて、読み出し処理の概要を説明する。読み出し処理は、ホストシステム10からアクセス要求された論理アドレスをキーとして、論理アドレス変換テーブル22を検索し、読み出し対象ブロックが格納されているストライプ番号とブロック番号を検索する(図12のステップS1202)。

【0051】次に、そのストライプ番号をキーとしてス

トライブ管理テーブル23を参照し、固定故障セクタ存在フラグを確認する(図12のステップS1204)。フラグがセットされている場合には、そのストライブは固定故障しているので、ホストシステム10に対してエラーを返す。セットされていない場合は、ストライブ番号とブロック番号を用いてディスク記憶装置をアクセスし、データを読み出してホストコンピュータへ送る(図12のステップS1206)。

【0052】もし、アクセス要求された論理アドレスに対応するブロックを含むストライブが、以前に故障ストライブとして検出されていれば、その内容が新ストライブに移動されている。その場合は、前述故障セクタが検出された際の処理の5)に於いて、論理アドレス変換テーブル22中の故障ストライブのストライブ番号を検索し、それを新ストライブのストライブ番号に変更しているため、新ストライブ中の該当するブロックをアクセスすることができる。

【0053】以上のように、上記実施形態によると、ログ書き込み方式記憶装置に於いて、代替セクタを使用しなくても、一部の領域の固定故障による障害を回避できる。従って、代替セクタを全て使い切った後でも、耐故障性を維持しつつ著しく性能が低下することなく、一部の領域の固定故障のためにディスク記憶装置を交換しなくても継続利用できる。また、ログ書き込み方式記憶装置を構成するディスク記憶装置に代替セクタ領域が確保されていない場合でも、あるセクタが固定故障した際にも継続して利用できる。更に、故障セクタの代わりに利用される領域は、通常はデータを記憶するために利用され、書き込み性能の向上に役立っている。即ち、セクタの固定故障に備えて全く使われない特別な領域を用意しておく必要もないので、ディスク記憶装置の利用効率もよい。

【0054】尚、本発明は上記実施形態に限定されるものではなく、種々の変形が可能なことは勿論である。

【0055】例えば上記実施形態に加え、故障ストライブの代わりに利用される領域の大きさを動的に制限する方法を説明する。先ず、図1に示すアクセス制限開始固定故障ストライブ数記憶手段26に、故障ストライブの代わりとして利用されるストライブ数の上限を設定しておく。この設定は、運用中に変更することも可能である。上記実施形態で示した故障セクタが検出された際の処理の最後に、以下の処理を追加する。

【0056】即ち、ストライブ管理テーブル23を検索し、故障ストライブ数を調べる(図13のステップS1302)。この結果と、アクセス制限開始固定故障ストライブ数記憶手段26中に記憶されている数とを比較し(ステップS1304)、故障ストライブの数が上回った場合は、以降はログ書き込み方式記憶装置に対する書き込みアクセスは禁止する(ステップS1306)。

【0057】本例によれば、故障ストライブ数がある一

定数以上に達した際には運用中止を即すことも可能となる。

【0058】ところで、上述実施形態にあっては、上記各種制御処理のプログラム情報をコンピュータ読取り可能な記憶媒体に格納し、当該記憶媒体を装着した装置が同プログラム情報を実行し、上述各種処理を行うようにしても良いものである。

【0059】

【発明の効果】以上詳記したように本発明によれば、ログ書き込み方式記憶装置にあって、代替セクタを全て使い切った場合であっても、耐故障性を維持しつつ著しく性能が低下することなく、一部の領域の固定故障のためにディスク記憶装置を交換しなくても継続して利用できる。また、ログ書き込み方式記憶装置を構成するディスク記憶装置に代替セクタ領域が確保されていない場合でも、あるセクタが固定故障した際にも継続して利用できる。更に、故障セクタの代わりに利用される領域は、通常はデータを記憶するために利用され、書き込み性能の向上に役立っている。即ち、セクタの固定故障に備えて全く使われない特別な領域を用意しておく必要もないので、ディスク記憶装置の利用効率もよい。その上、固定故障したセクタの代わりに利用される領域の大きさを動的に制限することも可能となるため、固定故障したセクタ数がある一定数以上に達した場合には運用中止を即すことも可能となる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係わるログ書き込み方式記憶装置の概略構成を示すブロック図。

【図2】同実施の形態に係わり、論理アドレス変換テーブルの概念的構成を示す図。

【図3】同実施の形態に係わり、ストライブ管理テーブルの概念的構成を示す図。

【図4】同実施の形態に係わり、ストライブとブロック関係を示す図。

【図5】同実施の形態に係わり、ストライブ中のブロックのディスク記憶装置への配置を示す図。

【図6】同実施の形態に係わり、ストライブの状態を示す図。

【図7】同実施の形態に係わり、ログ書き込み方式記憶装置の制御部の処理を示すフローチャート。

【図8】同実施の形態に係わり、論理アドレス変換テーブルの状態を示す図。

【図9】同実施の形態に係わり、ストライブ管理テーブルの状態を示す図。

【図10】同実施の形態に係わり、書き込み処理を示すフローチャート。

【図11】同実施の形態に係わり、有効ブロック数の更新処理を示すフローチャート。

【図12】同実施の形態に係わり、読み出し処理を示すフローチャート。

【図13】本発明の一実施形態の変形例を示すフローチャート。

【図14】従来のログ書き込み方式ディスク記憶装置の概略構成を示す図。

【符号の説明】

10…ホストシステム

20…ログ書き込み方式ディスク記憶装置20

21…制御部

22…論理アドレス変換テーブル

23…ストライプ管理テーブル

24…冗長情報生成手段

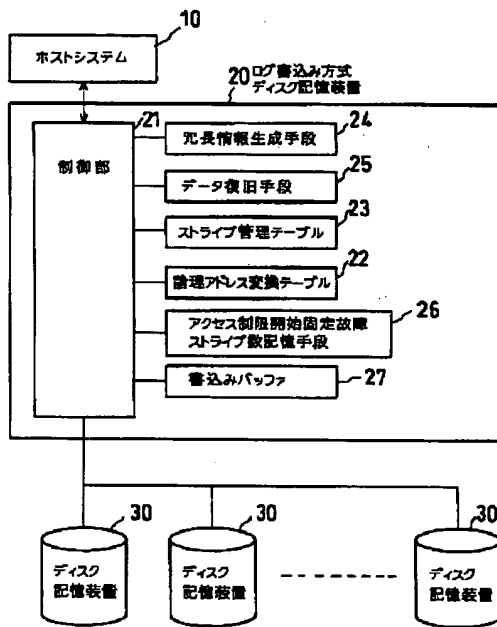
25…データ復旧手段

26…アクセス制限開始固定故障ストライプ数記憶手段

27…書き込みバッファ

30…ディスク記憶装置

【図1】



【図2】

論理アドレス	ストライプ番号	ブロック番号
L0		
L1		
L2		
⋮		

【図3】

ストライプ番号	有効ブロック数	固定故障セクタ存在フラグ
S0		
S1		
S2		
⋮		

【図4】

Si	B0	B1	B2	B3
	B4	B5	B6	B7

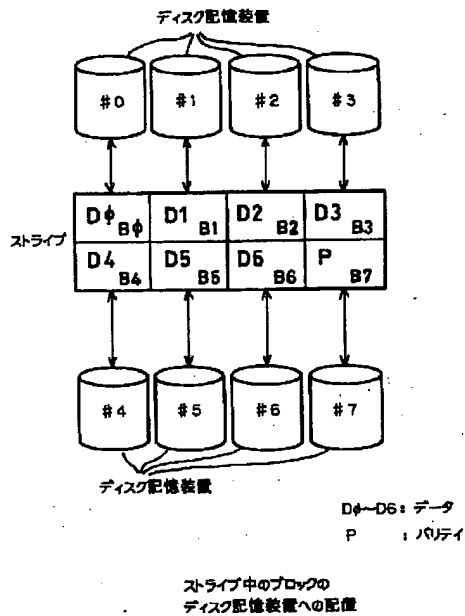
Si: ストライプ番号iのストライプ

B0~B7: ブロック番号

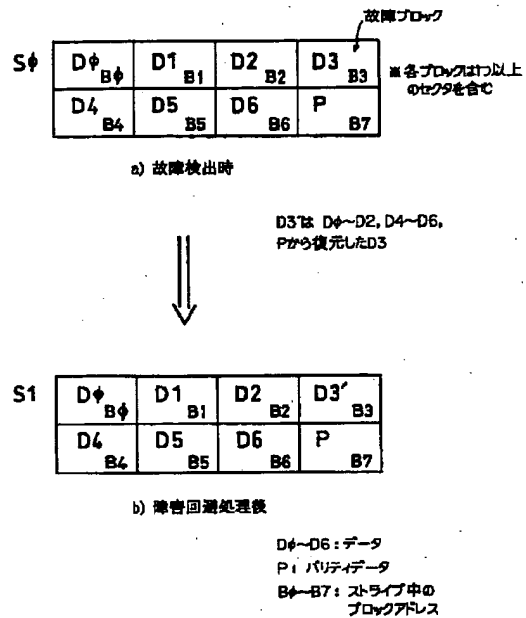
(ストライプ中のブロックのアドレス)

ブロック中にはデータか、データを復旧
するのに必要な冗長情報が置かれる

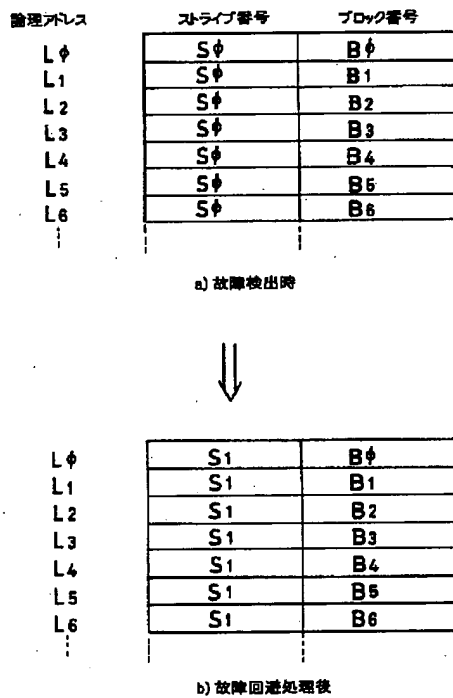
【図5】



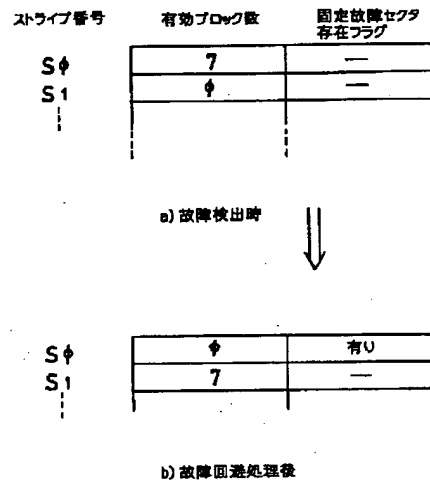
【図6】



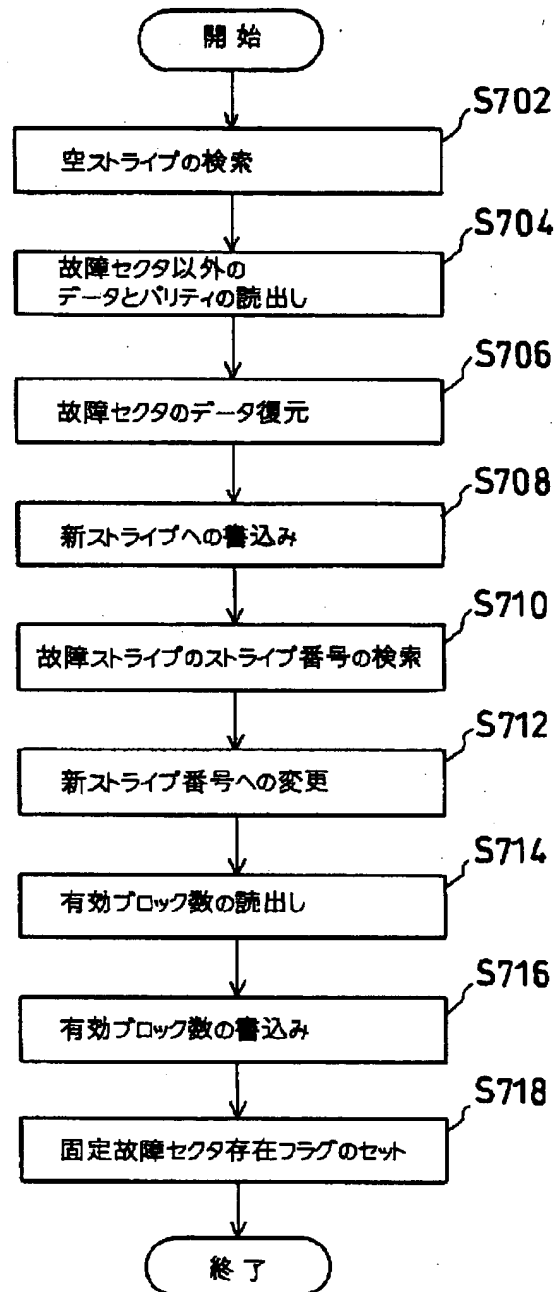
【図8】



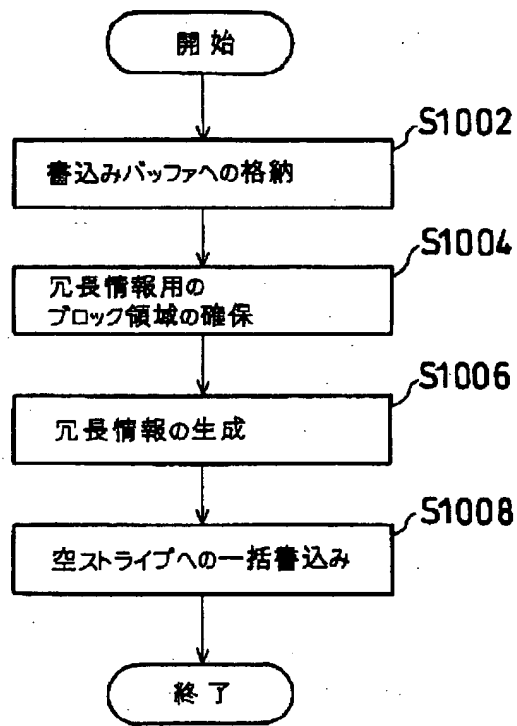
【図9】



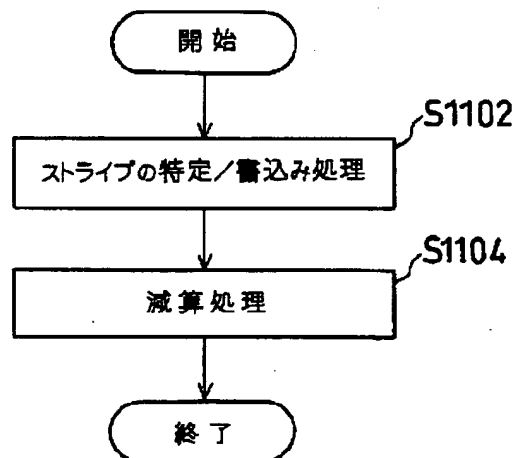
【図7】



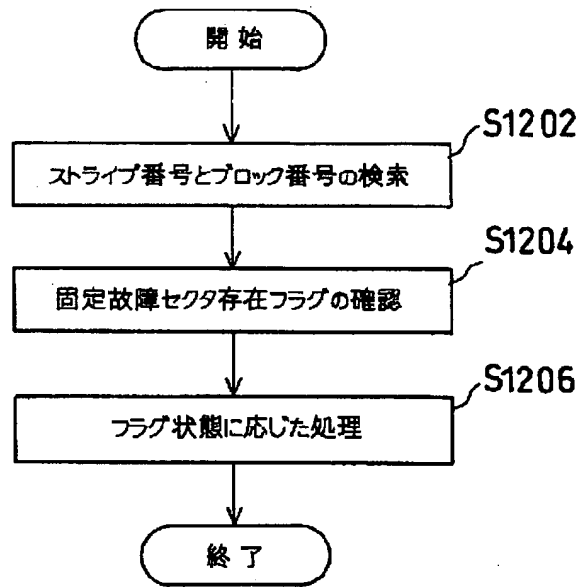
【図10】



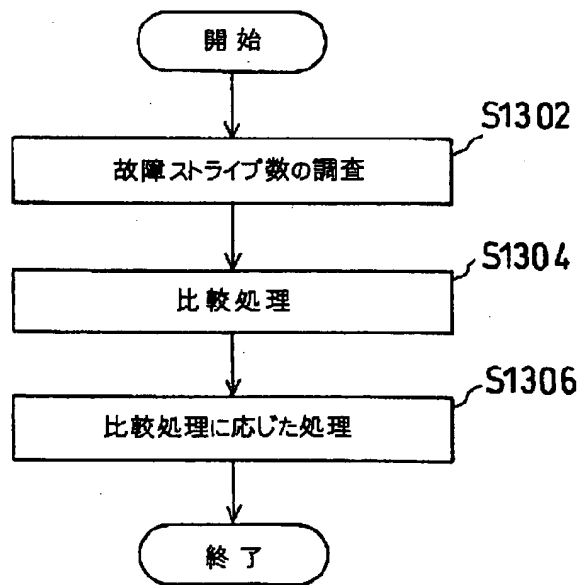
【図11】



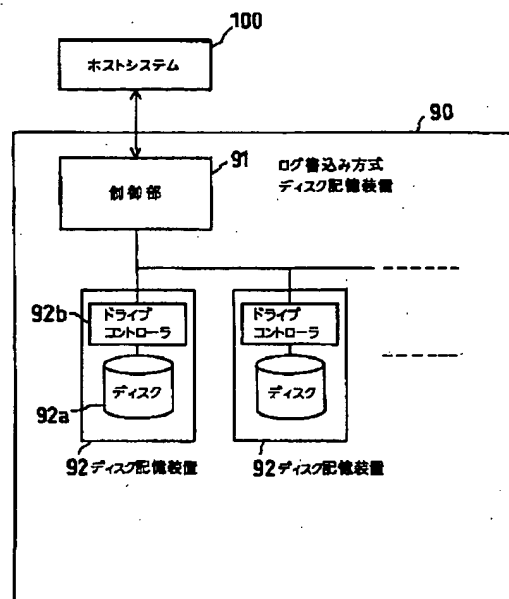
【図12】



【図13】



【図14】



THIS PAGE BLANK (USPTO)